

PHOTONE: EXPLORING MODAL SYNERGY IN PHOTOGRAPHIC IMAGES AND MUSIC

Niklas Rönnerberg

Linköping University
Media and Information Technology
SE-581 83 Linköping, Sweden
niklas.ronnerberg@liu.se

Jonas Löwgren

Linköping University
Media and Information Technology
SE-581 83 Linköping, Sweden
jonas.lowgren@liu.se

ABSTRACT

We present Photone, an interactive installation combining photographic images and musical sonification. An image is displayed, and a dynamic musical score is generated based on the overall color properties of the image and the color value of the pixel under the cursor. Hence, the music changes as the user moves the cursor. This simple approach turns out to have interesting experiential qualities in use. The composition of images and music invites the user to explore the combination of hues and textures, and musical sounds. We characterize the resulting experience in Photone as one of modal synergy where visual and auditory output combine holistically with the chosen interaction technique. This tentative finding is potentially relevant to further research in auditory displays and multimodal interaction.

1. INTRODUCTION

The starting point for us was the use of musical sonification to augment information visualization. With musical sounds we mean deliberately designed and composed sounds, based on a music-theoretical and aesthetic approach. Adding sound as a complementary modality to an image is related to the research area of sonification. One of the aims of sonification is to use sound to enhance and clarify visual representations of data, and to simplify the understanding of these [1, 2, 3, 4]. Sonification using musical sounds is interesting as the use of musical elements gives better control of the design of the sounds, and invokes potentially useful musical properties such as timbre, harmonics, melodies, rhythm, tempo, and amplitude (for more information about musical properties, see for example [5]).

Our previous applications of musical sonification include situation awareness in a monitoring/control task, where peripheral information on a radar screen for air traffic controllers was sonified to support the formation of a Gestalt of the airspace using musical sounds in different pitches and with different rhythms [6]. Sonification was also used for data exploration in scatter plots and parallel coordinates for discrimination of visual density levels using musical sounds of different pitch and timbre [7]. Furthermore, sonification was evaluated in relation to perception of color intensity [8] where three sonification conditions were used and interactively changed in relation to intensity levels in the visual repre-

sentations: cut-off frequency of a band-pass filter, amplitude level, and a combination of these two.

In all these applications, music served successfully as a complementary modality to enhance the performance of primarily visual tasks. The hierarchy between the visual and the auditory was clear. However, when we applied a similar approach to photographic images for hedonic experience in an interactive installation called Photone, we found the relationship between image and sound to be less clear-cut – the visual and the auditory appeared to be contributing equally to the multimodal experience.

Combining images and music is nothing new in itself. The history is full of excellent examples of composed music for images and motion pictures. There are also different approaches to visualizing music and for turning an image into music. In most cases, music is composed to images or generated from images based on their denotative and connotative meaning. Our artistic intention in Photone was another: By building the sonification upon pixel values of hue and brightness, that is, syntactic rather than semantic properties of an image, we aimed to cut through conventional ways of seeing to a more foundational level. The resulting experience of Photone cannot be understood as a simple sum of visual and auditory stimuli; our tentative analysis is that the fine-grained details of interaction form the glue between the visual and auditory modalities and the experience emerges as a synergistic whole. We propose the concept of *modal synergy* to capture this phenomenon. More on this below, but let us first try to convey a sense of the synergistic interaction experience we are talking about.

A short video demonstration of Photone to complement the following vignettes can be found here:
<https://vimeo.com/246964768>

2. VIGNETTES: USING PHOTONE

Consider a first example (see Figure 1). When locating the cursor in the dark center of the leftmost flower, the music is muffled and somewhat anticipatory in its low pitch and intensity. As you move the cursor slowly along one of the petals, the fine-grained textural changes invoke different notes forming a unique melody. Intensity and pitch increase with the brightness, but the music is still homogeneous in its harmony (the red harmony, as you are about to find out). Then, you cross the border between the petal and the orange background, and the music shifts to a different basic chord. The orange background has much less texture, which yields a more ambient and less articulated melodic quality. You move back and forth across the soft gradations and the more abrupt changes in the orange region, attentive to the differences, and perhaps try to find a rewarding rhythm by going back and forth across an intensity



This work is licensed under Creative Commons Attribution Non Commercial 4.0 International License. The full terms of the License are available at <http://creativecommons.org/licenses/by-nc/4.0>



Figure 1: Photo: Niklas Rönnberg.

shift. Moving on to exploring the blue, green and yellow areas of the image introduces new harmonics, giving you reason for further comparison and impromptu composition by going back and forth between the different color areas. This simple example already illustrates how image and music unfold together in an engaging interactive experience – the visual properties of the image suggest moves to explore the musical effects, and the musical qualities of the emerging composition suggest moves to exploit certain visual properties.



Figure 2: Photo: Niklas Rönnberg.

In the next example (see Figure 2), the lighted windows against the dark façade of the building on the left form distinct shifts in brightness, serving the same purpose here that the keys serve on a piano keyboard. The cursor is drawn there more or less involuntarily, and you move at different speeds along the rows of windows, concentrating on the resulting rhythm and its interplay with the ambient chords. From there, it is only a short step to playing around with the ripple textures on the surface of the stream, and the dynamic range between the lighted patches under the stone bridge and the dark pillars.

Going back and forth horizontally in the area above the glass (see Figure 3) provides a clear and engaging sense of how the red and the blue harmonics relate to each other, while each of them contain enough brightness variation to provide for interest-



Figure 3: Photo: Patric Ljung.

ing melodic possibilities in themselves.

The final example (see Figure 4) emphasizes the subtle combination of melodies and harmonics that lies in some images waiting to be discovered. The foliage has a fine-grained and sharp-edged texture yielding interesting melodic possibilities, and you may spend a long time there feeling how the autumn colors harmonize with each other. But at some point, you move the cursor from the foliage into the blue sky – and the triumphant quality of the resulting harmonic shift is quite remarkable and rewarding.



Figure 4: Photo: Rickard Englund.

What all these examples aim to show is how the interaction with photo and music has holistic qualities where it would be pointless to insist that one modality augments or supplements the other. What they also demonstrate is a specific way to look at images as collections of pixels with colors and brightness values, and the temporal trajectories formed by spatial movement across the surface of the image. We will return to these points in the concluding discussion, after a presentation of how Photone is designed and constructed.

3. DESIGN OF THE SONIFICATION

As previous studies (i.e. [6, 7, 8]) have suggested that amplitude, timbre, and pitch are useful musical elements to sonify visualiza-

tion of data, the composition in Photone was made with these elements as the basis. In the previous studies, data points were read as pixel values and the sonification was changed according to these values; the mapping between pixel values and musical elements was similar in Photone. The design and composition rationale for Photone was to compose the fundamental elements in a style inspired by electronic drone music, where the musical elements would be changed by the user's interaction and exploration.

An image in digital form consists of numerous small dots, pixels, where the color in each pixel is normally represented with three numbers, one each for the three color channels red, green, and blue (hence this technique for color reproduction is referred to as the RGB color space). Black is represented with no light (using the value 0) and white is represented with maximum light intensity (the value 255) in all three color channels simultaneously, and virtually any shade of color can be reproduced by adjusting the values in the different color channels [9].

The musical sonification in Photone starts with determining the predominant hue of a given image: red, green, blue, yellow, cyan, magenta, or white (see Figure 5). The musical expression differs in terms of harmonics and melodies between these hues. The intention is to create slightly different impressions of, for example, a whiter image compared to a greener one. The composition within each hue consists of six musical elements that are affected by the user interaction and the color intensity in the image (see Figure 6). The overall harmonic ambience is composed for each color channel (red, green, blue) to range from a simple two-tone chord to more complex chords, depending on the intensity and complexity of the color hues in the image. This creates an impression of a simpler and subtler harmonic ambience for darker and more basic colors, while brighter composite hues are sonified with more complex harmonic chords. Even if the harmonic and melodic components in Photone differ between a more red image and a more blue image, the overall musical expression is somewhat similar. Photone reads the pixel values in an image without regards to the depicted motive. The differences in musical expression lie rather in the user's interaction and exploration of the hues and textures in the images.

The three color channels (i.e. red, green, and blue) are related to different melodic components as well, creating different intervals and different melodic movements as the intensity in the color channels change. An increase in intensity generates a melody going upwards, and an intensity decrease yields a downwards-going melody. In areas with lower intensity, two low frequency bass tones are used to amplify the impression of darker shades of color. These bass tones are then attenuated as the intensity levels in the image increase to further enhance the feeling of different light levels in the image. A high-pitched light intensity chord with high frequency components is used to create an airy and high-intensity feeling in image areas with bright colors. For image areas containing pure white, a bell-like sound is played to accentuate the dazzling intensity of white. Similarly, for areas with very low intensity levels, a low frequency downwards-sweeping sound is played to emphasize the change in intensity from different shades of color to darkness.

4. IMPLEMENTATION

The implementation of the interactive sonification is done in SuperCollider 3.8.0, a programming environment for real-time audio synthesis [10, 11]. The RGB values of an input image are saved

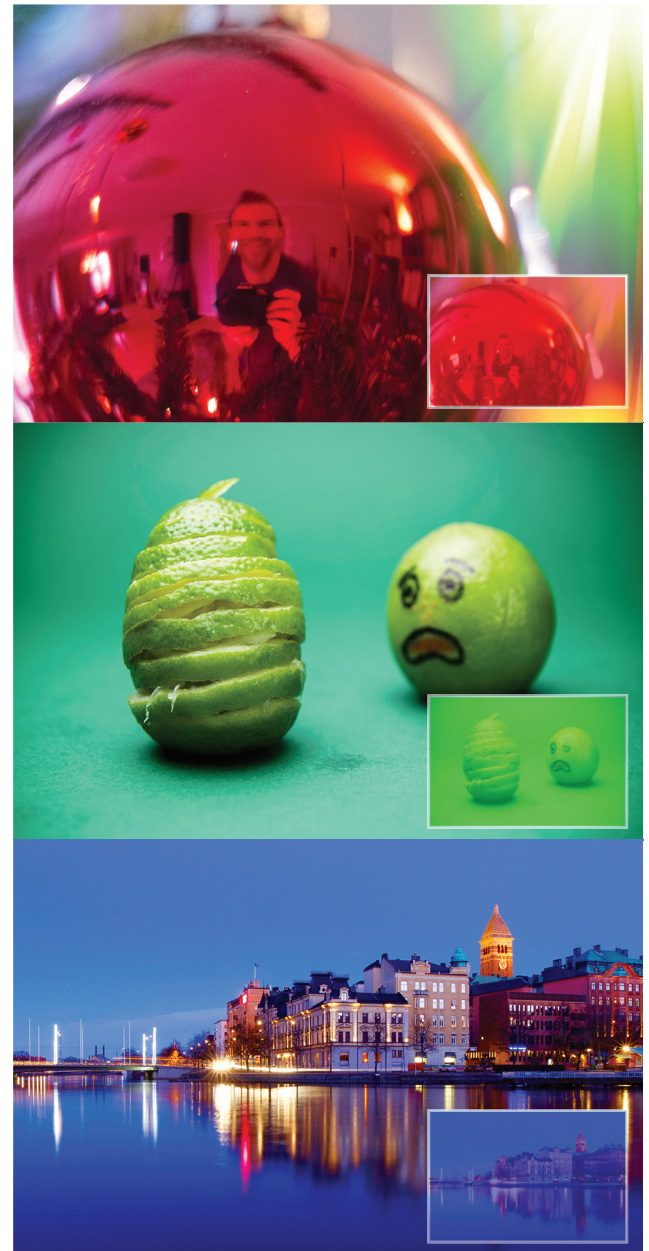


Figure 5: In Photone, the prominent hue of color of an image is determined, and the harmonics and melodic components selected accordingly. First (counted from above) a mainly red image, second a mainly green image, and third a mainly blue image. The small recessed images show the predominant hues. Photos (from above): Niklas Rönnberg, Patric Ljung, Paula Žitinski Elías.

from Matlab R2016a as text files that are read by SuperCollider for adjustment of the sound according to the intensity levels of the different color channels.

The composition consists of six musical elements (see Figure 6). These elements are 1) the overall harmonic ambience, 2) melodic components, 3) bass tones, 4) a high light intensity chord, 5) a bell-like sound for pure white, and 6) a low frequency sweep

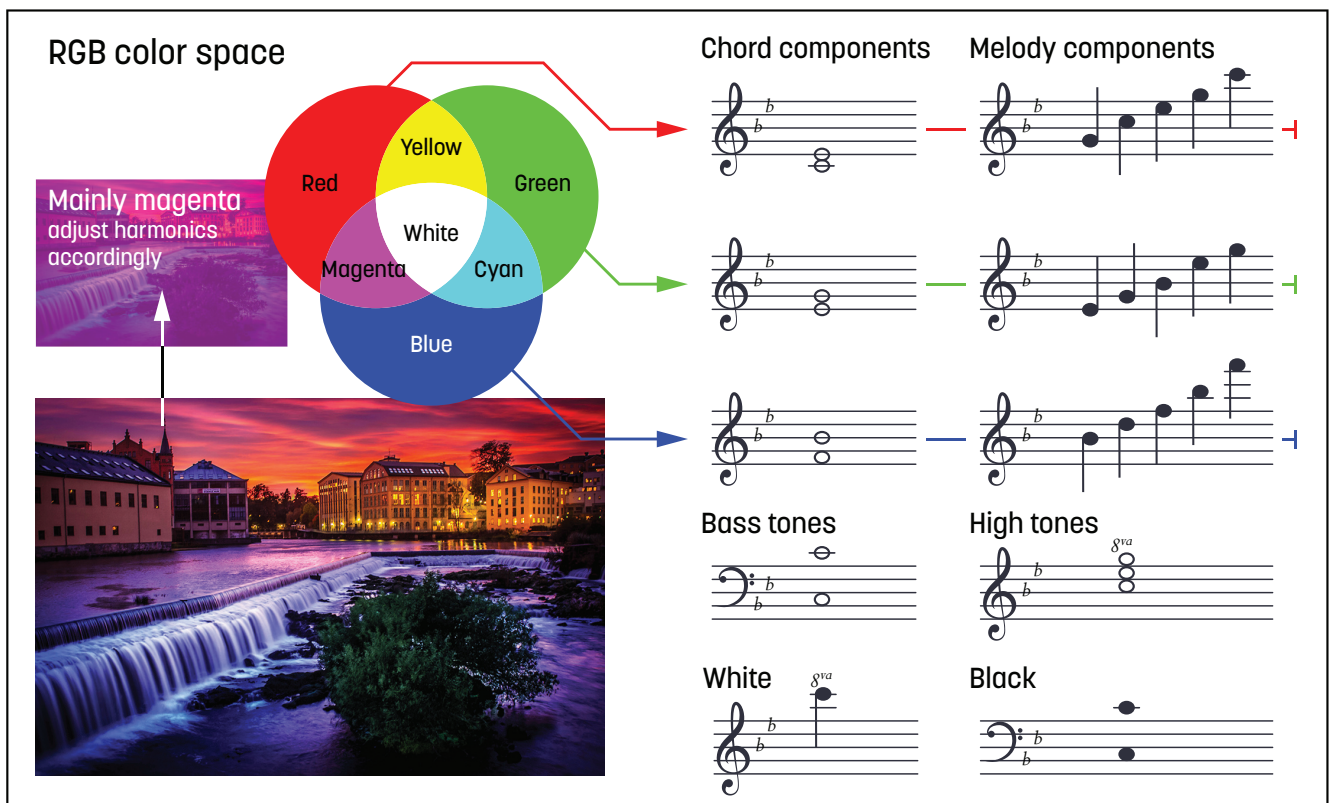


Figure 6: In Photone, each image is classified to a certain hue of color and the sonification is adjusted accordingly. The RGB color channels have different chords building up the ambient sound as well as different melodic components. The rest of the musical elements (the bass tones, the high chord, the bell-sound for white, and the synth sweep for black) are consistent within a certain main color shade, and are dependent on the intensity levels for each pixel. Photo: Paula Žitinski Elías.

for pure black. Of these musical elements, the overall harmonic ambience and the melodic components were composed for three color channels. Each image is determined to be mainly red, green, blue, yellow, cyan, magenta, or white and the composition used, in terms of harmonics and melodies, will differ between these colors.

The *harmonic ambience* consists of two-tone intervals multiplied over five octaves, creating a broad harmonic ambience with eleven tones for each color channel (see Figure 6). The red channel, for example, has the tones C2, Eb2, C3, Eb3, C4, Eb4, C5, Eb5, C6, Eb6, and C7. This creates a two-tone interval when the pixel under the cursor has light intensity only in one primary color channel (e.g. red, green, or blue), a four-tone interval when two color channels are used (e.g. red and green are blended together to create yellow, green and blue for cyan, or red and blue for magenta), and a six-tone interval when there is information in all three color channels simultaneously (e.g. white or bright shades of color). Each tone consists of seven triangle oscillators, where six oscillators are slightly detuned in relation to the fundamental frequency (-10, -7, -4, +4, +7, and +10 cents relative the fundamental frequency). The light intensity (0 to 255) of the individual color channel determines the sound level between 50% attenuation of the sound and no attenuation. If the intensity is 0 there is no sound for that color channel. The intensity for each color channel also controls the cut-off frequency of a second order band-pass filter (between 100 and 4000 Hz), through which the two-tone intervals

for each color channel pass. This results in a rich harmonic content ambient sound, where the harmonics varies in sound level and frequency components according to intensity levels in the three color channels independently of each other.

The *melodic components* consists of five tones (e.g. G5, C6, Eb6, G6, and C7 for the red channel), played one tone at a time, for each color channel (see Figure 6). The intensity level is divided into five steps and one of the tones is used accordingly. Each tone is built up with seven triangle waves in the same way as the harmonic ambience, but with a more pronounced attack and an attenuated sustain period. Also, similar to the harmonic ambience, the melodic components vary in amplitude level and in band-pass filter cut-off frequency. The *bass tones* consist of two tones (e.g. C2 and C3) (see Figure 6), that are only present when the overall intensity level is less than 200 of a total of 765, to fill up the ambient music when the harmonic ambience and the melodic components are attenuated. The *high light intensity chord* is composed with three tones (e.g. C7, Eb7, and G7), and played only when the overall light level is above 750. Both these instruments consist of three sawtooth oscillators each. The bass tones pass through a low-pass filter, while the high chord passes through a serial combination of a phaser, a flanger, and a resonant low-pass filter to create the impression of a more non-static changing sound. The *short bell-like sound* (e.g. the tone C8), used to sonify a white area in the image, consists of one triangle wave oscillator with a short

amplitude envelope. The *downwards sweeping sound* (e.g. C2 and C3), used to emphasize a dark region, consists of two sawtooth oscillators and a high resonance low-pass filter sweeping the cut-off frequency from high to low. Finally, all sounds are mixed together and passed through a reverb effect.

A simple screen saver is also implemented in the system, where a circle is moving in random directions across the screen and the sonification alters according to the pixel values under the center of the circle. The screen saver starts after 30 seconds of inactivity, and the photos are changed every fourth minute. The screen saver is used to draw attention to the Photone by displaying photos and creating an ambient sound.

5. OUTCOME AND IMPACT: NOT FOR EVERYONE

Photone was exhibited for a period of two weeks in a science center for the general public focusing on visualization and interactive experience. During this time, the science center had approximately 5000 visitors. Even if there is no exact count of the number of visitors exploring Photone, guides and staff at the science center estimate 100+ visitors. For these two weeks, the system was observed by guides and staff at the science center. Moreover, visitors had the ability to volunteer free-form text comments. There was no specific task the visitors were supposed to do in Photone. It was rather an open invitation to explore the cross-modal experience provided by Photone.

Our informal analysis of the data from the exhibition suggests that some visitors merely toyed with the interaction for a very brief period of time. At other occasions there were groups of visitors discussing the system while taking turns interacting with it, exploring the photos and the musical sounds. Then, there were sessions where visitors spent considerable time exploring the images and sounds, and the occasional comment indicating a visitor having a more or less transformative experience.

"First of all, [I] would like to say that the exhibition [i.e. Photone] was an interesting experience! Seems to be an exciting research area. On Saturday's visit, I expressed to my company that I would have liked to be able to sit for hours and explore this".

A tentative conclusion might be that Photone attracts only some of the visitors to a public exhibition space, but the visitors engaging more deeply in the interaction are drawn into a more or less captivating state of exploration similar to the vignettes outlined above. Our conjecture at this point is that engagement of this kind might correlate with musical aptitude and a reflective state of mind; more systematic studies are of course needed to validate any such speculations. It is also worth noting that Photone implements some of the established practices for facilitating exploration in a public space, such as the notion of progressive internal complexity and the creation of a frame for performative interaction, but not others such as engaging the full body and challenging situated norms [12]. This may also contribute to explaining the wide diversity in terms of visitors' engagement.

6. FUTURE DESIGN AND RESEARCH IDEAS

The tentative conclusions and feedback from users have given rise to some future design and research ideas. It would be interesting to further explore the interplay between colors, musical elements,

and the perception of emotions in a future version of Photone. Musical sounds are well adapted, at least on a more general level, to conveying meaning and emotions (see for example discussions in [13, 14]). Not only music affects emotions but also colors are associated with emotions [15], and there are correlations between the emotional associations of music and of colors [16]. On a more general level the interactive musical sonification used in Photone could adapt more to the prominent shade of color in each image. For example, it would be interesting to make a mainly red image sound distinctly different from a blue image, and with corresponding emotional impressions of the musical sounds as well as of the colors in the image. The harmonic ambience and the melodic components could be composed to instill the emotions in the user that are also present through the predominant color in the image. Apart from obvious aspects of musical sounds, such as major chords and melodies for more positive color schemes and minor chords and melodies for more negative colors, there are other interesting musical elements to consider. One example is melodic movements where an ascending pitch is generally perceived as more positive, while a descending pitch is perceived as more negative [17]. This could be utilized in the melodic components to further emphasize the impression of the overall color hue. Other musical elements that could be fruitful to explore is the harmonic timbre (harmony versus dissonance), as more complex harmonic timbres (i.e. a simple major or minor chord, such as C or Cb, versus a suspended or a diminished chord, such as Csus4 or Cdim) are more captivating for a listener compared to simpler harmonic timbres [18]. A truly dissonant combination of tones is also experienced as more unpleasant compared to harmonious chords [19]. Both the complexity of the harmonics in chords and the use of dissonant tones in a chord could be used to further link the musical sounds, the experience of color, and the emotional response.

The overall lightness of the image could also be used to create the impression that the musical sounds are even more connected to and derived from the image. For example, lighter and brighter colors are associated with higher pitched tones [20, 21, 22]. There might also be a relation between the experienced difference between high and low pitch tones, and the experienced difference between brighter and darker areas (see for example discussions in [23, 24, 25]). The overall brightness in the image could be used to select different ranges in the harmonic ambience with higher pitches for brighter images, and lower pitches for darker images. The use of a more soft or dull timbre will also be experienced as more negative compared to a brighter timbre [17]. In a similar way, the melodic components could be transposed in accordance with the brightness level in the image.

In the current version of Photone the amplitude and the timbre change (by changing the cut-off frequency of a band-pass filter) depending on the intensity levels in the three color channels. This affects each color channel separately, but in a future version of Photone the overall brightness could also affect the overall amplitude and frequency content in all of the musical sounds. Interesting to note is that a louder sound is more activating and engaging for the listener/user than a quieter sound [18]. The perception of loudness is also mapped to brightness, for example in a image, via the amplitude of a sound [26]. Consequently, a brighter image would have louder musical sounds with higher pitched harmonics compared to a generally darker image.

Further more, it would be fascinating to use image processing algorithms to detect patterns and shapes in the images. The pattern recognition could detect different levels of patterns, such as con-

trasts between and differences in hues of color and brightness. The level of patterns detected could then be used to adjust the tempo of some musical elements, for example the tempo of an arpeggio or of a percussive rhythmic sound, where a more cluttered image would yield musical elements with a faster tempo. The shape recognition could be used to detect different shapes or movements in the image, such as downwards going shapes that could lead to a descent in pitch when the user explores that shape, or the level of roughness in the image which in turn could be mapped to the cut-off frequency of a filter or the wave-shaping (e.g. the pulse width of a square wave sound) of the musical sounds leading to more harsh sounds when the roughness in an area of the image is high, and more smooth sounds when the roughness is low.

Finally, Photone could also be expanded to a completely immersive experience, with stereoscopic images and spatial sounds, using VR or AR techniques. It would also be possible to involve a tactile modality, in order to explore the potential of multimodal synergy.

We believe that these future design and research ideas would make Photone into an even more interesting interactive experience of photographic images and musical sounds. The next implementation of Photone will have logs of all mouse cursor movements with time stamps. This data will be used to identify interaction patterns and areas of interest. These data will be accompanied by data collected from an observational study, as well as interviews about the user experience. More systematic studies of the interactions in Photone, and about the interplay between music and image, would also lead to new and interesting insights in modal synergies.

7. DISCUSSION: TOWARDS MODAL SYNERGY

The work we report here is explorative and to some degree tentative. Still, it has already yielded insights of some significance to the field of auditory/multimodal display and, more generally, to interaction design.

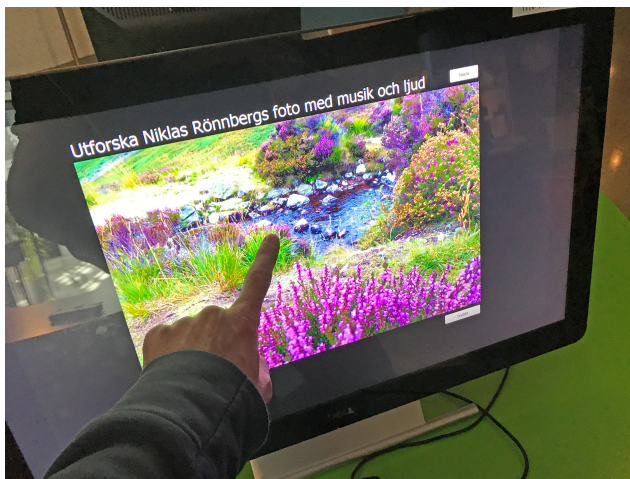


Figure 7: Photone can be explored either with a computer mouse or a touch screen with quite different experiences of the interaction. Photo: Niklas Rönnberg.

We have experimented with Photone in two different hardware configurations, i.e., one with mouse and cursor input and one with

a vertical touchscreen (see Figure 7). It is quite striking how different the two conditions are with respect to the resulting experience of use. When using a mouse and cursor, the experiential state is one of prolonged and incisive exploration (looking closely, manipulating precisely). The touchscreen, on the other hand, becomes more of a way to sample the general musical timbre of different color and brightness areas. This difference is mainly due to the current implementation of the sonification algorithm, where the sound modulation is based on the exact pixel value under the cursor. Another part of the reason is that music unfolds over time, and moving the cursor slowly over a textured area in the image while listening attentively to changes in the music is simply more comfortable with a mouse that allows the hand to rest. This issue is solvable as such, of course, with a persistent finger-controlled pixel-level cursor offset from the point of touch, but the point is rather this: *the fine-grained, bodily details of the interaction must be considered part of the use experience together with the visual and auditory output modalities*. Working with a conventional vertical touchscreen, for example, would properly speaking motivate redesigning the sonification towards a more expressive interaction with large, sweeping movements and "broad strokes". Furthermore, the touchscreen adds an interesting feature to the interaction, which is the possibility to go from one pixel in a certain area of the image to another pixel in a different area. This enables, as mentioned above, sampling of different color and brightness values and the musical sounds derived from these, without the transitions between these areas. Thus, the touchscreen permits a nonlinear exploration, and to some degree facilitates the use of the image as a musical instrument when "playing the image".

This line of reasoning aims to show how image, sound and interaction are tightly integrated in Photone. In fact, we propose to see the three elements as aspects of the same emergent experience in use. Multimodality and multimodal interaction have been established concepts in interaction design for a long time, but it is notable that the concepts have not been unpacked systematically from an experiential point of view to any significant extent. Most accounts of multimodality seem to focus on how to use different input modalities (keyboard, mouse, touch, speech, etc.) in concert and how different output modalities supplement each other [27]. In that sense, the prevalent conceptual understanding of multimodality seems to be more or less on par with where we started in musical sonification: "How can sonification supplement an existing visual modality to enhance visual presentations and increase the users' task performance?".

Our current work aims to go beyond this supplementary approach to explore how the visual and auditory modalities come together and form an experiential whole through interaction. We have focused specifically on photographic image and musical sound, and it is worth noting that our approach to the images is rather deconstructive. Musical sonification of photographic images based on their denotative and connotative meanings ("this is a picture of a flower" and "this picture suggests tropical mystique", respectively) would essentially amount to the kind of impression management found in movie music composition, or more simply in the approach known as auditory icons (such as the sound of a bee buzzing around the flower and the sound of a gong and wind chimes for tropical mystique). In Photone, the sonification is driven by pixel values of hue and brightness, that is, syntactic rather than semantic properties of an image. Other syntactic image features that could be used in the same way include edges, shapes and spatial frequency measures, see discussion above. Our aim

was to cut through conventional ways of looking at a photographic image. At the risk of sounding pretentious, we would argue that the current combination of photographic images, sonification algorithms, and mouse-and-cursor input represents the potential for a more foundational way of seeing – similar to how a beginning portrait artist needs to unlearn ideas of noses, ears and facial expressions in order to see lines, shapes and hues. To us, this represents one example of how image, sound and interaction together form something more than a mere addition of supplementary modalities. We call this emergent phenomenon *modal synergy*, and we propose that further design-led inquiries in this direction may be a way to add to our constructive knowledge of auditory displays and multimodal interaction.

8. REFERENCES

- [1] T. Hermann, A. Hunt, and J. G. Neuhoff, *The Sonification Handbook*, 1st ed. Berlin, Germany: Logos Publishing House, 2011.
- [2] T. Pinch and K. Bijsterveld, *The Oxford Handbook of Sound Studies*. Oxford University Press, 2011.
- [3] K. Franinovic and S. Serafin, *Sonic Interaction Design*. Cambridge, MA, USA: MIT Press, 2013.
- [4] L. Philipsen and R. S. Kjærgaard, *The Aesthetics of Scientific Data Representation: More than Pretty Pictures*. Denmark, Europe: Routledge. Routledge Advances in Art and Visual Studies., 2018.
- [5] S. J. Deliège, I., *Perception and Cognition of Music*. Hove, East Sussex: Psychology Press Ltd., 1997.
- [6] N. Rönnerberg, J. Lundberg, and J. Löwgren, “Sonifying the periphery: Supporting the formation of gestalt in air traffic control,” in *The 5th Interactive Sonification Workshop*, ser. ISON-2016. Germany: CITEC, Bielefeld University, 2016, pp. 23–27.
- [7] N. Rönnerberg and J. Johansson, “Interactive sonification for visual dense data displays,” in *The 5th Interactive Sonification Workshop*, ser. ISON-2016. Germany: CITEC, Bielefeld University, 2016, pp. 63–67.
- [8] N. Rönnerberg, “Sonification enhances perception of color intensity,” in *Proceedings of IEEE VIS 2017, InfoVIS posters*, ser. IEEE VIS 2017, 2017.
- [9] C. Ware, *Information Visualization: Perception for Design*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2013.
- [10] J. McCartney, “Supercollider: A new real-time synthesis language,” in *The International Computer Music Conference*, ser. ICMC’96. New York, NY, USA: ACM, 1996, pp. 257–258.
- [11] —, “Rethinking the computer music language: Supercollider,” *Computer Music Journal*, vol. 26, no. 4, pp. 61–68, 2002.
- [12] M. Hobye, “Designing for homo explorans : open social play in performative frames,” Ph.D. dissertation, Faculty of Culture and Society Malmö University, Malmö, Sweden, 2014. [Online]. Available: <http://muep.mau.se/handle/2043/16510>
- [13] N. Rönnerberg and J. Löwgren, “The sound challenge to visualization design research,” in *Proceedings of EmoVis 2016, ACM IUI 2016 Workshop on Emotion and Visualization*, ser. EmoVis 2016. Sweden: Linköping Electronic Conference Proceedings, 2016, pp. 31–34.
- [14] F. J. L. L. Tsuchiya, T., “Data-to-music api: Real-time data-agnostic sonification with musical structure models,” in *Proc. 21st International Conference on Auditory Display (ICAD 2015)*, 2006, pp. 244–251.
- [15] A. N. Gilbert, A. J. Fridlund, and L. A. Lucchina, “The color of emotion: A metric for implicit color associations,” *Food Quality and Preference*, vol. 52, pp. 203–210, 2016.
- [16] S. E. Palmer, K. B. Schloss, Z. Xu, and L. R. Prado-Leon, “Musiccolor associations are mediated by emotion,” *PNAS*, vol. 110, pp. 203–210, 2013.
- [17] P. Juslin and P. Laukka, “Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening,” *Journal of New Music Research*, vol. 33, pp. 217–238, 2004.
- [18] S. A. Iakovides, V. T. Iliadou, V. T. Bizeli, S. G. Kaprinis, K. N. Fountoulakis, and G. S. Kaprinis, “Psychophysiology and psychoacoustics of music: Perception of complex sound in normal subjects and psychiatric patients,” *Annals of General Hospital Psychiatry*, vol. 3, pp. 1–4, 2004.
- [19] K. J. Pallesen, E. Brattico, C. Bailey, A. Korvenoja, J. Koivisto, A. Gjedde, and S. Carlson, “Emotion processing of major, minor, and dissonant chords: A functional magnetic resonance imaging study,” *Annals New York Academy of Sciences*, vol. 1060, pp. 450–453, 2005.
- [20] W. G. Collier and T. L. Hubbard, “Musical scales and brightness evaluations: Effects of pitch, direction, and scale mode,” *Musicae Scientiae*, vol. 8, pp. 151–173, 2004.
- [21] L. E. Marks, “On cross-modal similarity: Auditoryvisual interactions in speeded discrimination,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 13, pp. 384–394, 1987.
- [22] J. Ward, B. Huckstep, and E. Tsakanikos, “Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all?” *Cortex*, vol. 42, pp. 264–280, 2006.
- [23] J. Best, *Colour Design: Theories and Applications*, 2nd ed. Duxford, United Kingdom: Elsevier Ltd. : Woodhead Publishing, 2017.
- [24] R. Bresin, “What is the color of that music performance?” in *Proc. International Computer Music Conference ICMC 2005*. San Francisco, CA: International Computer Music Association, 2005, pp. 367–370.
- [25] S. E. Palmer, T. A. Langlois, and K. B. 1 Schloss, “Music-to-color associations of single-line piano melodies in non-synesthetes,” *Multisensory Research*, vol. 29, pp. 157–193, 2016.
- [26] R. W. Pridmore, “Music and color: Relations in the psychophysical perspective,” *Color Research & Application*, vol. 17, pp. 57–61, 1992.
- [27] M. A. Mathias Nordvall, “Perception, meaning and transmodal design,” in *Proceedings of Design Research Society 50th Anniversary Conference*, ser. DRS 2016, 2016, pp. 1–12.